Protein Science (1999), *8*:2806–2812. Cambridge University Press. Printed in the USA.

FOR THE RECORD

Mass spectral characterization of a protein–nucleic acid photocrosslink

MACE C. GOLDEN,[1,2,3] KATHERYN A. RESING,[1] BRIAN D. COLLINS,[2]
MICHAEL C. WILLIS,[2] AND TAD H. KOCH[1]

[1]Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215
[2]Nexstar Division of Gilead Sciences, Inc., Boulder, Colorado 80301
[3]Department of Chemistry, U.S. Air Force Academy, Colorado Springs, Colorado 80840

(RECEIVED March 15, 1999; ACCEPTED September 17, 1999)

Abstract: A photocrosslink between basic fibroblast growth factor ($bFGF_{155}$) and a high affinity ssDNA oligonucleotide was characterized by positive ion electrospray ionization mass spectrometry (ESIMS). The DNA was a 61-mer oligonucleotide photoaptamer bearing seven bromodeoxyuridines, identified by in vitro selection. Specific photocrosslinking of the protein to the oligonucleotide was achieved by 308 nm XeCl excimer laser excitation. The crosslinked protein–nucleic acid complex was proteolyzed with trypsin. The resulting peptide crosslink was purified by PAGE, eluted, and digested by snake venom phosphodiesterase/alkaline phosphatase. Comparison of the oligonucleotide vs. the degraded peptide crosslink by high performance liquid chromatography coupled to an electrospray ionization triple quadrupole mass spectrometer showed a single ion unique to the crosslinked material. Sequencing by collision induced dissociation (MS/MS) on a triple quadrupole mass spectrometer revealed that this ion was the nonapeptide TGQYKLGSK (residues 130–138) crosslinked to a dinucleotide at Tyr133. The MS/MS spectrum indicated sequential fragmentation of the oligonucleotide to uracil covalently attached to the nonapeptide followed by fragmentation of the peptide bonds. Tyr133 is located within the heparin binding pocket, suggesting that the in vitro selection targeted this negative ion binding region of $bFGF_{155}$.

Keywords: bromodeoxyuridine; mass spectrometry; photocrosslinking; sequencing

Crosslinking of proteins to associated nucleic acids followed by sequencing is a useful method for learning structural information about the protein–nucleic acid interface, information frequently unavailable from crystallography or NMR spectroscopy. When these more precise structural tools are applicable, crosslinking information can also facilitate interpretation of the raw data toward solving the structure (Horvath et al., 1998). Crosslinking is accomplished both chemically and photochemically. Photochemical crosslinking methods often have the advantage of causing minimal perturbation of the protein–nucleic acid interface (Meisenheimer & Koch, 1997). An important method involves the substitution of 5-bromouridine for uridine in RNA (Tanner et al., 1988; Gott et al., 1991; Stump & Hall, 1995) or 5-bromo-2′-deoxyuridine for thymidine in DNA (Allen et al., 1991; Hicke et al., 1994). The 5-bromouracil chromophore specifically crosslinks to aromatic and sulfur bearing amino acid residues in close proximity when excited in the region of 310 nm (Dietz & Koch, 1987; Dietz & Koch, 1989; Norris et al., 1997). The XeCl excimer laser, emitting monochromatic light at 308 nm, is particularly effective for crosslinking proteins to nucleic acids bearing this substitution (Meisenheimer et al., 1999). Relevant to the results described here, model studies of the photoreaction of 5-bromouracil and 5-bromouridine with N-acetyltyrosine N-ethylamide have established that covalent bond formation occurs between the 5-position of the nucleoside base and the position ortho to the hydroxyl group of tyrosine with loss of HBr as shown in Figure 1 (Dietz & Koch, 1987; Norris et al., 1997).

Peptide sequencing is most commonly performed by Edman degradation of enzymatic digests of the crosslinked protein–nucleic acid complex. Mass spectrometry has emerged as a complementary and/or superior technique for the sequencing. The sequencing of proteins crosslinked to oligonucleotides poses a special problem because sequencing of peptides is done in the positive ion mode, while sequencing of nucleic acids is normally done in the negative ion mode. However, several research groups have shown that matrix assisted laser desorption ionization (MALDI) and HPLC/electrospray ionization (ESI) mass spectrometry (MS) can be used in the positive mode to establish the molecular mass of proteins chemically and photochemically crosslinked to oligonucleotides, as well as the peptide-oligonucleotide crosslink generated by proteolysis (Jensen et al., 1993; Bennett et al., 1994; Jensen et al., 1994; Connor et al., 1998b; Wong et al., 1998). Sequencing by collision-induced dissociation (CID) of model pep-

Reprint requests to: Tad H. Koch, Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215; e-mail: tad.koch@colorado.edu.

Abbreviations: bFGF, basic fibroblast growth factor; CID, collision-induced dissociation; DTT, dithiothreitol; ESI, electrospray ionization; HPLC, high performance liquid chromatography; HSA, human serum albumin; LC/MS, HPLC coupled to MS; MALDI, matrix assisted laser desorption ionization; MS, mass spectrometry; MS/MS, tandem mass spectrometry; PAGE, polyacrylamide gel electrophoresis; PBS, phosphate buffered saline; SELEX, selective evolution of ligands by exponential enrichment; ssDNA, single stranded DNA.

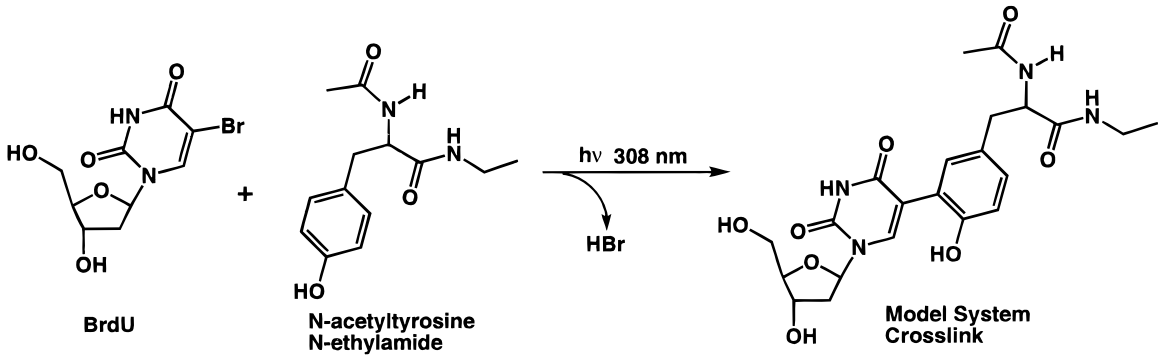

Fig. 1. Model photoreaction between 5-bromouridine (BrU) and N-acetyltyrosine N-ethylamide showing the nature of the covalent bond.

tides chemically crosslinked to a homogeneous oligonucleotide T_6 and of peptides photochemically crosslinked to mono or dinucleotides has been demonstrated (Jensen et al., 1996; Connor et al., 1998a). A crosslink has also been characterized with MALDI mass spectrometry coupled with N-terminal microsequencing and alkaline RNA hydrolysis (Urlaub et al., 1997). Although these cases show the potential for mass spectral characterization of protein–nucleic acid crosslinks, MS sequencing of a crosslinked high affinity protein–nucleic acid complex by CID has not been previously reported.

The utility of mass spectral identification of a protein–nucleic acid photocrosslink is now demonstrated using basic fibroblast growth factor $bFGF_{155}$ crosslinked to a high affinity single stranded DNA aptamer as a model system. bFGF is a member of a family of polypeptides that are regulators of cell proliferation, differentiation, and motility (Baird & Klagsbrun, 1991; Vlodavsky et al., 1991). It is released from cells and subsequent binding to cell surface heparin facilitates binding to the bFGF receptor. Prevention of heparin binding affects cell growth and differentiation (Rapraeger et al., 1991). A 61-mer deoxyoligonucleotide aptamer that binds to bFGF was selected from a ssDNA library using a variation of an in vitro selection methodology referred to as SELEX (Tuerk & Gold, 1990). The variation (photoSELEX) uses 5-bromo-2′-deoxyuridine in place of thymidine in the library and selects for both affinity and high yield crosslinking, producing a photoaptamer (Jensen et al., 1995; M.C. Golden, B.D. Collins, M.C. Willis, & T.H. Koch, in prep.). This study has established that the site of crosslinking of the photoaptamer to bFGF occurs in the heparin binding pocket. This was achieved by HPLC/ESI positive ion MS/MS on a snake venom phosphodiesterase/alkaline phosphatase digest of the tryptic peptide bearing the crosslink.

Results: The macromolecules photocrosslinked in this study were the 155 amino acid variant of bFGF and its high affinity 61-mer single stranded deoxyoligonucleotide photoaptamer shown in Figure 2. The 61-mer bears seven 5-bromo-2′-deoxyuridines and was synthesized as double stranded DNA from a synthetic ssDNA template using the polymerase chain reaction (PCR). The sense strand was partitioned from the antisense strand by denaturing polyacrylamide gel electrophoresis (PAGE), because migration of the anti-

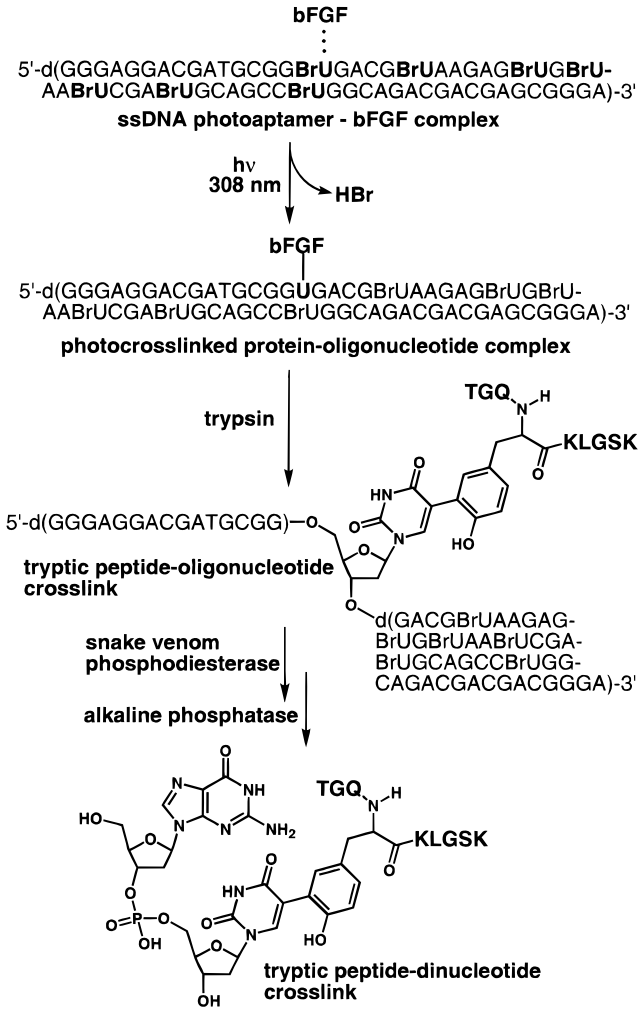

Fig. 2. Photochemical crosslinking of a photoaptamer to bFGF and subsequent enzymatic cleavage with trypsin followed by snake venom phosphodiesterase/alkaline phosphatase to prepare a sample for mass spectral analysis. Snake venom phosphodiesterase is proposed to leave the deoxyguanosine on the 5′-side of the oligonucleotide because of steric interference from the crosslink.

sense strand was slowed by the presence of two biotin groups covalently bound to its 3′-end. The sense strand (850 pmol), partially 5′-end labeled with ^{32}P, was photocrosslinked to bFGF by irradiation at 308 nm with a XeCl excimer laser.

The crosslinked protein-oligonucleotide was enzymatically degraded to simplify the structure for mass spectral characterization of the crosslink. First, the crosslinked protein-oligonucleotide was digested with trypsin, and the crosslinked peptide-oligonucleotide was separated from other tryptic peptides and unreacted DNA by PAGE to yield 150 pmol of product (54% recovery from 280 pmol of bFGF). Of this quantity, 60 pmol was further degraded for mass spectral sequencing, and 60 pmol was used for Edman sequencing to confirm the mass spectral result. To simplify the mass spectral analysis, most of the nucleic acid sequence was next removed from the crosslinked peptide-oligonucleotide using snake venom phosphodiesterase/alkaline phosphatase (Fig. 2). The sample was then applied to a C18 reverse phase capillary HPLC column directly coupled to the inlet of the mass spectrometer (LC/MS) to separate the resulting peptide–nucleotide crosslink from the enzymes and other contaminants. Of the 60 pmol allotted for the mass spectral analysis, 20 pmol was used to compare with a parallel sample of the oligonucleotide digested in the same manner as the crosslinked material. The crosslinked sample had one high intensity ion not present in the noncrosslinked sample (Fig. 3A). This unique ion was doubly charged with m/z 768.8 ($MH^+ = 1{,}536.6$ Da). The charge state was determined by the presence of ions bearing sodium ion (779.9) or potassium (788.4) in place of a proton. No singly or triply charged ion for this peptide was observed.

With the remaining 40 pmol, an on-line LC/MS/MS spectrum (Fig. 3B,C) from collision-induced dissociation of the m/z 769 ion was obtained. The MS/MS spectrum shows high intensity fragment ions at 1,287.8/644.2 and 1,091.6/546.4 Da, corresponding to singly and doubly charged forms resulting from loss of dehydrodeoxyguanosine (249.6 Da, predicted 249.1 Da) and phosphodehydrodexoyribose (196.3 Da, predicted 196.0 Da) as shown in Figure 4A. Observation of a small ion of mass 151.6 Da, corresponding to the predicted MH^+ for guanine, confirmed that one of the nucleotides was guanosine. A series of y_1-y_5 ions corresponding to the sequence KLGSK and weak b_2 and b_3 ions identify the peptide as the incomplete proteolytic product TGQYKLSGSK (predicted unmodified $MH^+ = 1{,}069.2$ Da). The difference between the observed mass and the predicted mass for the unmodified peptide indicates that the peptide had two nucleotides at the crosslink. As one of these was deoxyguanosine, the mass difference indicates that the other had to be deoxyuridine, as expected from the predicted photochemistry (Fig. 1). The crosslink itself is identified at the tyrosine, because both the $y_6 + U$ ion (806 Da) and the $b_4 + U$ ion (560 Da) are observed along with b_3 and b_2 ions (Figs. 3C, 4B). Thus the crosslink is confirmed by two separate ion products. The Edman sequencing of a parallel sample gave the same peptide sequence, TGQXKLGS(K?), but also showed a low yield sequence derived from incomplete digestion at the preceding arginine residue (Arg129). This low yield peptide was not observed in the mass chromatogram.

Discussion: These results demonstrate that a single crosslink was generated by photocrosslinking the photoaptamer to bFGF and provide important insights into the properties of the crosslinked peptide. The mass of the tryptic peptide was consistent with the expected photochemistry of crosslinking, confirming the mecha-

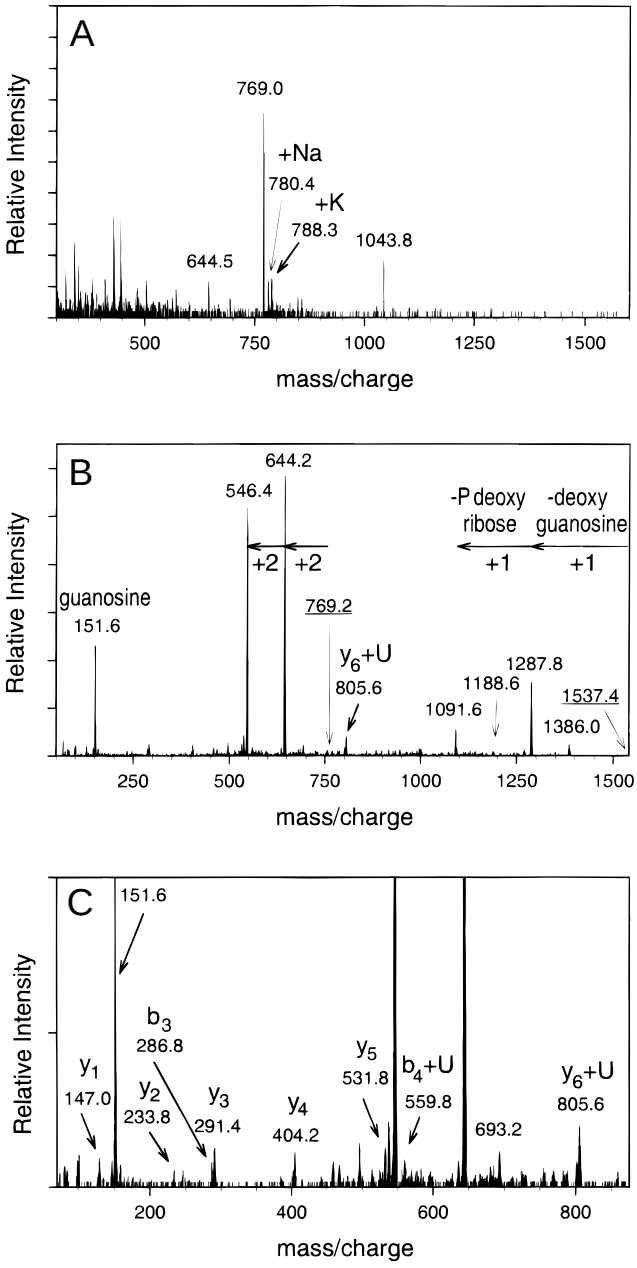

Fig. 3. Mass spectrum and tandem mass spectrum (MS/MS) of the tryptic peptide crosslinked to the dinucleotide shown in Figure 2. Proposed ion structures are shown in Figure 4. **A:** Mass spectrum derived from summing three scans of the LC/MS of the nuclease treated, gel purified tryptic peptide from the cross-linked sample. The ion at m/z 644 is a fragment ion of the 769.2 Da peptide (see **B**). The ions smaller than m/z 510 Da are also present in the control and are primarily due to solvent contaminants. The ion at 1,043.8 Da could not be accounted for, but did not appear to be a peptide ion. **B:** Full MS/MS spectrum of the doubly charged parent ion of 769.2 Da showing the assignment of ions from fragmentation of the dinucleotide portion of the crosslink. The major fragment ions were the singly and doubly charged ions resulting from loss of dehydrodeoxyguanosine followed by dehydrodeoxyribosyl phosphate. **C:** Expanded view of the MS/MS spectrum in **B**, showing assignment of ions from fragmentation of the peptide portion of the crosslink. Fragment ions for y_1 through y_5 and $y_6 + U$ were observed, as well as b_2 (158.6 Da, not labeled), b_3, and $b_4 + U$. The ions at 693.2 and 1386.0 Da (in **B**) are the doubly and singly charged ions from loss of guanine. Also evident are the doubly charged $y_6 + U$ and $y_7 + U$ and its dehydration product (495.8, 467.2, and 459.2 Da, not labeled).

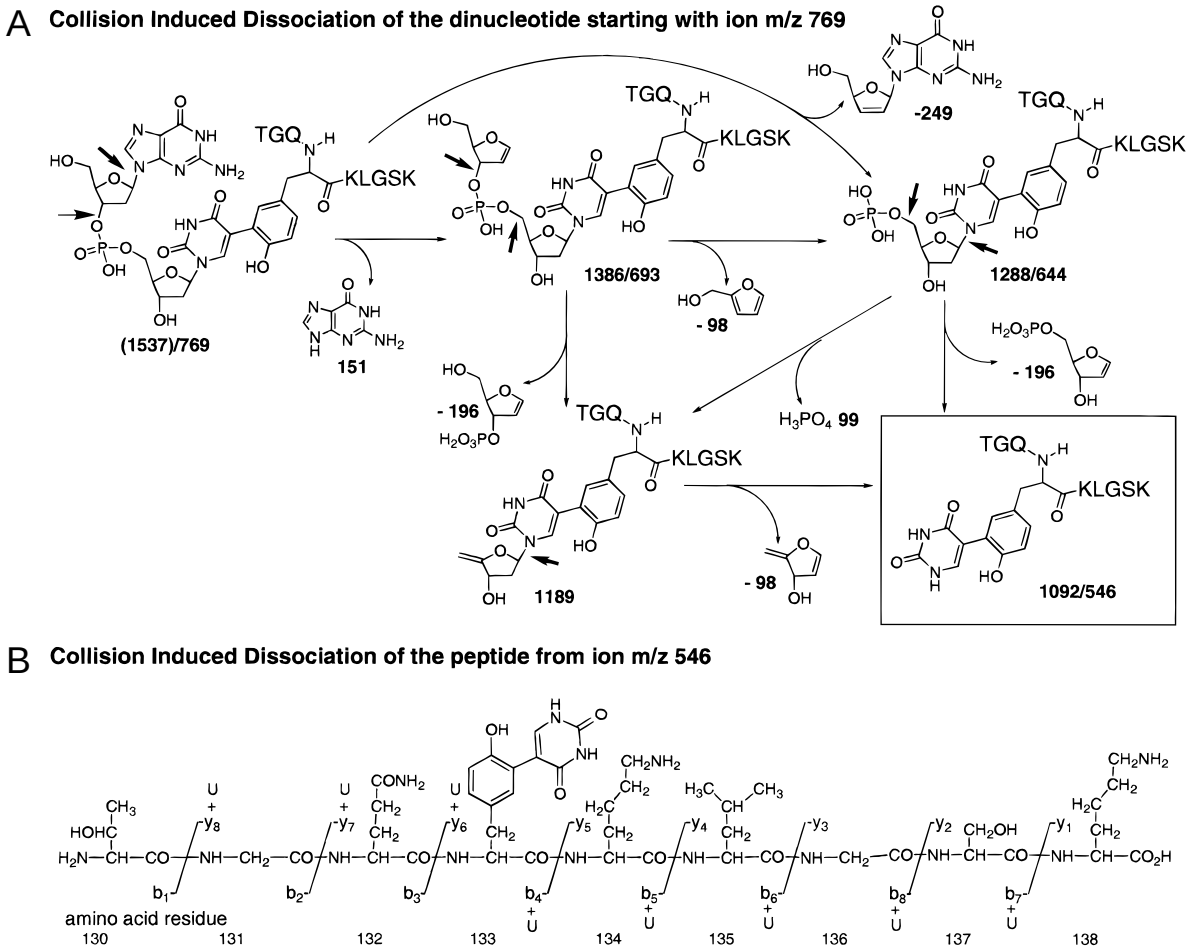

Fig. 4. Proposed ion structures for the fragmentation of the tryptic peptide crosslinked to the dinucleotide. **A:** Initial fragmentation of the nucleotide portion. The heavy black arrows indicate fragmentation sites. **B:** Subsequent fragmentation of the peptide portion to give *b* and *y* ions (Biemann, 1990a, 1990b).

nism determined with model compounds. Both trypsin and snake venom phosphodiesterase did not digest normally susceptible bonds adjacent to the crosslink. In the MS/MS spectrum, the molecular ion appeared to fragment initially by loss of components of the dinucleotide in various combinations (Fig. 4A). The mechanism is likely an elimination reaction analogous to that observed in negative ion MS/MS sequencing of oligonucleotides (McLuckey & Habibi-Goudarzi, 1993). After the molecular ion degraded to uracil crosslinked to the tyrosine residue of the nonapeptide, the predominant pathway for further degradation gave low intensity *b* and *y* ions (Figs. 3C, 4B) (Biemann, 1990a, 1990b), which established the crosslink at the tyrosine residue. The stability of the bond from uracil to tyrosine with respect to further degradation is consistent with a bond from the 5-position of the uracil ring to the aromatic ring of the tyrosine as predicted from the model reaction shown in Figure 1. The data suggest that most of the collision energy went into cleavage of the nucleotide, so that fragment ions resulting from cleavage of the peptide bonds were of low abundance. This might present difficulties where sample is limiting; however, the facile cleavage to the peptidyl uracil fragment ion indicates that use of multistage fragmentation methods, for example, in an ion trap or with the high voltage orifice method, would provide sequence information with limiting sample and/or larger peptides.

The snake venom phosphodiesterase/alkaline phosphatase cleavage of the oligonucleotide portion of the crosslinked molecule to a tryptic peptide-dinucleotide crosslink simplified the mass spectrum. A consequence, however, was a lack of information about the location of the crosslink on the oligonucleotide. From the MS/MS, the dinucleotide could be either 5′-d(GU)-3′ or 5′-d(UG)-3′. We assume that it is the 5′-d(GU)-3′, because we expect that the crosslink would have a larger effect on the hydrolysis of the 5′-phosphodiester linkage by the snake venom phosphodiesterase, but it is also possible that both are present. This problem is unique to a crosslink between a high affinity protein and a photoaptamer bearing multiple bromodeoxyuridines or bromouridines. More commonly, photocrosslinking is achieved with oligonucleotides bearing a single halonucleotide substitution. In this particular case, where five of the six bromouridines had adjacent deoxyguanosines, the nucleotide crosslink at 5′-d(GUG)-3′ was established by Maxam–Gilbert sequencing (M.C. Golden, B.D. Collins, M.C. Willis, & T.H. Koch, in prep.). Use of other cleavage methods that would leave a larger piece of the oligonucleotide would be a useful adjunct to the simultaneous identification of the nucleotide and amino acid residues involved in the crosslink.

Mass spectral determination of the crosslinking site establishes a contact point for the photoaptamer relative to the bFGF heparin

binding site. Tyr133 is within the heparin binding site of bFGF which includes residues 128–138 (Moy et al., 1996) as shown in Figure 5. This site has several basic amino acid residues (Lys128, Arg129, Lys134, and Lys138) that are important for binding heparin. The basic amino acid residues within the heparin binding site plus Arg90, which lies adjacent to the photocrosslinking site Tyr133, are predicted to be important for the affinity binding of the negatively charged 61-nucleotide photoaptamer.

In summary, we have demonstrated a facile methodology for the isolation and characterization of a protein–nucleic acid photocrosslink, which involves protein as well as oligonucleotide enzymatic degradation followed by HPLC electrospray mass spectral analysis with collision-induced dissociation. The example illustrated here is the crosslink between bFGF and a high affinity ssDNA photoaptamer bearing bromodeoxyuridine nucleotides. Although relatively large amounts of material were available in this study, recent improvements in mass spectrometry instruments and low flow sample application methods (Wilm et al., 1996) should allow a significant increase in sensitivity.

Materials and methods: *Materials:* Recombinant human basic fibroblast growth factor ($bFGF_{155}$) expressed in *Escherichia coli* was purchased from Bachem California (Torrence, California), lot number SN 433. The lyophilized protein was resuspended in 1x phosphate buffered saline (PBS), 2 mM $MgCl_2$, and 0.01% human serum albumin (HSA) to a concentration of 10 μM. Aliquots (20 μL) were stored at −20 °C until ready for use. Snake venom phosophodiesterase I, calf alkaline phosphatase, dATP, dGTP, and dCTP were purchased from Pharmacia Biotech (Piscataway, New Jersey), trypsin and 5-bromo-2′-deoxyuridine triphosphate (dBrUTP) from Sigma Chemical Co. (St. Louis, Missouri), [γ^{32}P]ATP and T4 polynucleotide kinase from NEN Life Science Products (Boston, Massachusetts), *Taq* polymerase from Perkin Elmer (Foster City, California), and *Pwo* polymerase from Boehringer-Mannheim (Indianapolis, Indiana).

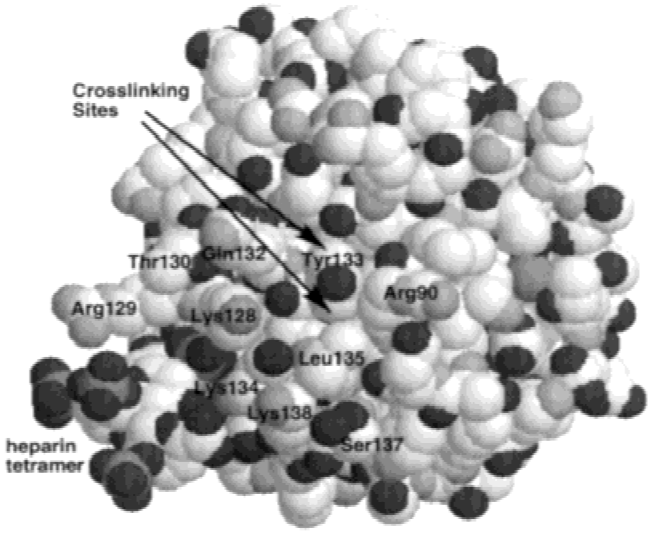

Fig. 5. A space filling model of bFGF bound to heparin tetramer (Faham et al., 1996) showing the ortho positions of the crosslinking amino acid residue Tyr133, created with RasMol from Brookhaven Protein Data Base file 1BFB. Other amino acid residues of the nonapeptide from tryptic digestion of the crosslinked protein–nucleic complex (Fig. 2) are also labeled as well as some additional basic amino acid residues that are likely involved in the binding of the deoxyoligonucleotide to the protein. The numbering of the residues is based on that for the 155 amino acid protein beginning with Met. In file 1BFB Tyr133 is Tyr125. Atoms are coded in gray scale with the order of increasing darkness being C, N, S, O.

Synthesis of 5-bromo-2′-deoxyuridine-substituted oligonucleotide: The 61-mer oligonucleotide photoaptamer was synthesized using the PCR to amplify a ssDNA template. The PCR was conducted in 1.5 mL with 30 pmol of template, 200 μM dATP, dGTP, dCTP, and dBrUTP, 18.75 units of *Taq* polymerase and 18.75 units of *Pwo* polymerase in a standard *Taq* PCR buffer (Sambrook et al., 1989) including 15 mM $MgCl_2$. To maximize amplification with the modified dNTP, all reagents other than polymerase enzymes were incubated at 75 °C for 3 min prior to amplification; then, 16 cycles of amplification (95 °C/40 s, 55 °C/60s, 72 °C/120 s) were performed using a Mini Cycler (MJ Research, Watertown, Massachusetts). The resulting dsDNA was concentrated with Centri-Sep 30 KMWCO filters (Schleicher and Schuell, Keene, New Hampshire) by centrifugation at 12,000 rpm for 10 min. Standard formamide gel loading buffer (45 μL) (Sambrook et al., 1989) including 0.1% sodium dodecyl sulfate (SDS) was used to resuspend concentrated DNA from the filter surface. To obtain sense ssDNA, the dsDNA was heated at 95 °C for 2 min and partitioned on a 12% polyacrylamide gel incorporating 7 M urea. The migration of the antisense strand was slowed due to the presence of two biotin groups covalently attached to the 3′-primer used in the PCR. The sense ssDNA strand was located on the gel by UV shadowing and subsequently excised. The gel band was crushed in 300 μL of 2 mM EDTA using a 1 cc syringe plunger until a homogeneous slurry was obtained. The slurry was vortexed at ambient temperature for 5 h and then centrifuged through a Centri-Sep 45 μm Costar filter at 14,000 rpm for 10 min. ssDNA was finally concentrated by standard NaOAc/ethanol precipitation and quantified by UV spectroscopy. A small portion of ssDNA (2 pmol) was 5′-radiolabeled with T4 polynucleotide kinase with subsequent gel purification as described above, except visualization was accomplished with a 1.5 min exposure of the gel on Kodak X-Omat film.

Protein–nucleic acid photocrosslinking: The oligonucleotide (850 pmol) and 5×10^4 cpm of 5′-^{32}P-labeled oligonucleotide was diluted to 500 μL with crosslinking buffer (1x PBS, 2 mM $MgCl_2$, 0.01% HSA, 1 mM dithiothreitol (DTT)) to give a 1.7 μM solution with 1,000 cpm/μL. bFGF (10 μM stock solution) was diluted to 500 μL with crosslinking buffer to a give a 0.567 μM solution. These two solutions were combined and incubated at 37 °C for 15 min. The solution was then transferred to a 1 mL poly methyl acrylate cuvette possessing a 1 cm path length and the cuvette placed in a 37 °C cuvette holder. The cuvette was irradiated with 1,500 pulses of 308 nm light from a Lumonix Model EX 700 XeCl excimer laser operating at 175 mJ per pulse. The laser intensity was attenuated by placing the cuvette 50 cm behind a quartz convex lens with a focal length of 10 cm, and with this configuration the sample received ~11 mJ/pulse of 308 nm irradiation.

Trypsin digestion and recovery of crosslinked peptide: Trypsin (400 μL of 5 mg/mL suspended in 0.1 M pH 8 Tris buffer) was added to the 1 mL solution of crosslinked protein–nucleic acid

and incubated overnight at 37 °C. The reaction mixture was divided into three portions and each was extracted with phenol/chloroform to remove nonnucleic acid constituents. The crosslinked peptide–nucleic acid in each portion was precipitated with NaOAc/ethanol, collected by centrifugation, and dissolved in formamide gel loading buffer. After heating at 95 °C for 2 min, the crosslinked peptide–nucleic acid was purified on a 40 cm 14% polyacrylamide gel with 7 M urea. A significant band migrating slightly slower than free nucleic acid was identified by 16 h exposure on Kodak Bio-Max film at −70 °C and subsequently excised. The gel slice was crushed to achieve a homogeneous slurry in 400 μL of deionized water, and the slurry was vortexed for 16 h. The crosslinked peptide–nucleic acid was recovered by centrifugation of the slurry through a 0.45 μm Centri-Sep Costar filter (14,000 rpm for 15 min) followed by NaOAc/ethanol precipitation and resuspension in 100 μL of deionized H_2O. The amount of recovered crosslinked peptide–nucleic acid was estimated to be 152 pmol by activity measurements in scintillation fluid compared with the total activity of the starting nucleic acid. A portion of the sample (60 pmol) was submitted to Midwest Analytical, Inc. (St. Louis, Missouri) for automated Edman degradation sequencing.

Snake venom phosphodiesterase/alkaline phosphatase digestion: Lyophilized snake venom phosphodiesterase I was resuspended to a concentration of 1 mg/mL with dilution buffer (110 mM Tris/pH 8.9, 110 mM NaCl, 15 mM $MgCl_2$, 50% glycerol in H_2O). Crosslinked peptide–nucleic acid (60 pmol) was diluted to a volume of 43.3 μL with deionized H_2O, and 0.8 μL of 1 mM $MgCl_2$, 3.5 μL of 0.5 M pH 7.5 Tris buffer, and 2.4 μL of resuspended SVP were added to give a total reaction volume of 50 μL. The reaction was then incubated overnight at 37 °C. Snake venom phosphodiesterase was deactivated by heating at 95 °C for 2 min. One Phor All Buffer (10x, 5.8 μL) and 1:20 diluted alkaline phosphatase (2 μL) were added with another incubation at 37 °C for 30 min. Finally, alkaline phosphatase was deactivated by heating at 95 °C for 2 min. As a control for mass spectral analysis, 60 pmol of nucleic acid (without the crosslink to bFGF) was subjected to identical digestion conditions.

Mass spectral analysis: Mass spectrometry (MS) and collision-induced dissociation (MS/MS) were accomplished with an API-III$^+$ triple-quadrupole mass spectrometer (Sciex, Norwalk, Connecticut) equipped with a nebulization assisted electrospray ion (ESI) source and a high-pressure collision cell. Samples were analyzed by HPLC using a 500 μm i.d. column packed manually with 300 Å Columbus C18 resin (Phenomenix, Torrance, California), equilibrated with 0.1% formic acid eluting at 40 μL/min and directly coupled to the mass spectrometer electrospray interface. The peptide-oligonucleotide crosslink was eluted from the column with a 2%/min gradient with acetonitrile at 20 μL/min, with the orifice voltage at 75 V, scanning at 0.6 s/scan. The 60 pmol of peptide-dinucleotide crosslink was used for two injections: the first 20 pmol for mass scans from *m/z* 300–1,600 Da with elution from 0–40% acetonitrile for determining the retention time and the second 40 pmol for sequencing the peptide-dinucleotide crosslink. The identification of the HPLC peak associated with the crosslinked peptide was facilitated by comparison of the chromatogram with that from injection of the control bearing compounds resulting from reaction of snake venom phosphodiesterase and alkaline phosphatase with the oligonucleotide.

Acknowledgments: This work was supported in part by grants from the Council for Tobacco Research and Nexstar Pharmaceuticals to THK, from the NIH (AR43768) to KAR, and by the Howard Hughes Medical Institute to Natalie Ahn for the purchase and maintenance of mass spectral equipment.

References

Allen T, Wick K, Matthews K. 1991. Identification of amino acids in lac repressor protein cross-linked to operator DNA specifically substituted with bromodeoxyuridine. *J Biol Chem 266*:6113–6119.

Baird A, Klagsbrun M. 1991. The fibroblast growth factor family. *Cancer Cells 3*:239–243.

Bennett SE, Jensen ON, Barofsky DF, Mosbaugh DW. 1994. UV-catalyzed cross-linking of *Escherichia coli* uracil-DNA glycosylase to DNA. *J Biol Chem 269*:21870–21879.

Biemann K. 1990a. Nomenclature for peptide fragment ions. *Methods Enzymol 193*:886–887.

Biemann K. 1990b. Sequencing of peptides by tandem mass spectrometry and high-energy collision-induced dissociation. *Methods Enzymol 193*:455–479.

Connor DA, Falick AM, Shetlar MD. 1998a. UV light-induced cross-linking of nucleosides, nucleotides and a dinucleotide to the carboxy-terminal heptad repeat peptide of RNA polymerase II as studied by mass spectrometry. *Photochem Photobiol 68*:1–8.

Connor DA, Falick AM, Young MC, Shetlar MD. 1998b. Probing the binding region of the single-stranded DNA-binding domain of rat DNA polymerase β using nanosecond-pulse laser-induced cross-linking and mass spectrometry. *Photochem Photobiol 68*:299–308.

Dietz TM, Koch TH. 1987. Photochemical coupling of 5-bromouracil to tryptophan, tyrosine and histidine, peptide-like derivatives in aqueous fluid solution. *Photochem Photobiol 46*:971–978.

Dietz TM, Koch TH. 1989. Photochemical reduction of 5-bromouracil by cysteine derivatives and coupling of 5-bromouracil to cystine derivatives. *Photochem Photobiol 49*:121–129.

Faham S, Hileman RE, Fromm JR, Linhardt RJ, Rees DC. 1996. Heparin structure and interactions with basic fibroblast growth factor. *Science 271*:1116–1120.

Gott JM, M CW, Koch TH, Uhlenbeck OC. 1991. A specific UV-induced RNA-protein crosslink using 5-bromouracil-substituted RNA. *Biochemistry 30*:6290–6295.

Hicke BJ, Willis MC, Koch TH, Cech TR. 1994. Telomeric protein-DNA point contacts identified by photocrosslinking using 5-bromodeoxyuridine. *Biochemistry 33*:3364–3373.

Horvath MP, Schweiker VL, Bevilacqua JM, Ruggles JA, Schultz SC. 1998. Crystal structure of the Oxytricha nova telomere end binding protein complexed with single strand DNA. *Cell 95*:963–974.

Jensen KB, Atkinson BL, Willis MC, Koch TH, Gold L. 1995. Using in vitro selection to direct the covalent attachment of human immunodeficiency virus type 1 Rev protein to high-affinity RNA ligands. *Proc Natl Acad Sci USA 92*:12220–12224.

Jensen ON, Barofsky DF, Young MC, vonHippel PH, Swenson S, Seifried SE. 1993. Direct observation of UV-crosslinked protein–nucleic acid complexes by matrix-assisted laser desorption ionization mass spectrometry. *Rapid Commun Mass Spectrom 7*:496–501.

Jensen ON, Barofsky DF, Young MC, vonHippel PH, Swenson S, Seifried SE. 1994. Mass spectrometric protocol for the analysis of UV-crosslinked protein-nucleic acid complexes. In: *Techniques in protein chemistry V*. New York: Academic Press. pp 27–37.

Jensen ON, Kulkarni S, Aldrich JV, Barofsky DF. 1996. Characterization of peptide-oligonucleotide heteroconjugates by mass spectrometry. *Nucleic Acids Res 24*:3866–3872.

McLuckey SA, Habibi-Goudarzi S. 1993. Decomposition of multiply charged oligonucleotide anions. *J Am Chem Soc 115*:12085–12095.

Meisenheimer KM, Koch TH. 1997. Photocross-linking of nucleic acids to associated proteins. *Crit Rev Biochem Mol Biol 32*:101–140.

Meisenheimer KM, Meisenheimer PL, Koch TH. 1999. Nucleoprotein photocrosslinking using halopyrimidine-substituted RNAs. *Methods Enzymol.* Forthcoming.

Moy FJ, Seddon AP, Böhlen P, Powers R. 1996. High-resolution solution structure of basic fibroblast growth factor determined by multidimensional heteronuclear magnetic resonance spectroscopy. *Biochemistry 35*:13552–13561.

Norris CL, Meisenheimer KM, Koch TH. 1997. Mechanistic studies relevant to bromouridine enhanced nulcleoprotein photocrosslinking. Possible involvement of an excited tyrosine residue of the protein. *Photochem Photobiol 65*:201–207.

Rapraeger AC, Krufka A, Olwin B. 1991. Requirement of Heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science 252*:1705–1708.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning*. Plainview, New York: Cold Spring Harbor Laboratory Press.

Stump WT, Hall KB. 1995. Crosslinking of an iodo-uridine-RNA hairpin to a single site on the human U1A N-terminal RNA binding domain. *RNA 1*:55–63.

Tanner NK, Hanna MM, Abelson J. 1988. Binding interactions between yeast tRNA ligase and a precursor transfer ribonucleic acid containing two photoreactive uridine analogues. *Biochemistry 27*:8852–8861.

Tuerk C, Gold L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science 249*:505–510.

Urlaub H, Thiede B, Müller E-C, Brimacombe R, Wittmann-Liebold B. 1997. Identification and sequence analysis of contact sites between ribosomal proteins and rRNA in *Escherichia coli* 30 S subunits by a new approach using matrix-assisted laser desorption/ionization mass spectrometry combined with N-terminal microsequencing. *J Biol Chem 272*:14547–14555.

Vlodavsky I, Bar-Shavit R, Ishai-Michaeli R, Bashkin P, Fuks Z. 1991. Extracellular sequestration and release of fibroblast growth factor: a regulatory mechanism? *Trends Biochem Sci 16*:268–271.

Wilm M, Shevchenko A, Houthaeve T, Brelt S, Schweigerer L, Fotsis T, Mann M. 1996. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature 379*:466–469.

Wong DL, Pavlovich JG, Reich NO. 1998. Electrospray ionization mass spectrometric characterization of photocrosslinked DNA-EcoR1 DNA methyltransferase complexes. *Nucleic Acids Res 26*:645–649.